FRONTIERS

in COMPUTING
+ MATHEMATICAL
SCIENCES

JANUARY 7
AND JANUARY 14 2019

Caltech

FRONTIERS IN COMPUTING + MATHEMATICAL SCIENCES 2019

ALL TALKS ARE ONE HOUR LONG, AND WILL BE HELD IN ANNENBERG 105. COFFEE BREAKS WILL BE HELD IN THE ANNENBERG LOBBY.

JANUARY 7

Time	Talk	Speaker
8:45am	Introduction	Andrew Stuart — Caltech
9:00am	Classical Verification of Quantum Computations	Urmila Mahadev — UC Berkeley
10:15am	Statistical Estimation Under Algebraic Constraints: Total Positivity	Elina Robeva — MIT
11:30am	Discrete Differential Geometry and Its Applications	Andrew Sageman-Furnas — Technical University of Berlin
2:00pm	The Cost of Crushing: Curvature-Driven Wrinkling of Thin Elastic Shells	Ian Tobasco — University of Michigan
3:15pm	Large-Scale Optimal Transport: Statistics and Computation	Jonathan Weed — MIT

JANUARY 14

| Time	Talk	Speaker
8:45am	Introduction	Andrew Stuart — Caltech
9:00am	Modeling Cell State Dynamics from Single-Cell Data and Its Implications in Cancer Development and Drug Resistance	Heyrim Cho — University of Maryland
10:15am	Extreme Events: Dynamics, Prediction and Mitigation	Mohammad Farazmand — MIT
11:30am	Solving Linear Equations	Rasmus Kyng — Harvard University
2:00pm	A Modern Maximum-Likelihood Approach for High-Dimensional Logistic Regression	Pragya Sur — Stanford University
3:15pm	An Aggregative Game Framework for the Analysis of Socio-Technical Systems	Francesca Parise — MIT

Modeling Cell State Dynamics from Single-Cell Data and Its Implications in Cancer Development and Drug Resistance

Heyrim Cho
University of Maryland, College Park

Recent advances in single-cell gene sequencing data and high-dimensional data analysis techniques are bringing in new opportunities in modeling biological systems with continuous phenotypic structured models. In this talk, we discuss ideas to develop continuum cell state models using high-dimensional single-cell gene data. In particular, dimension reduction techniques are applied to find the trajectories of hematopoietic stem cell states in the reduced differentiation space, then modeled as directed and random movement on the abstracted graph with PDEs. We simulate normal differentiation and abnormal processes of acute myeloid leukemia (AML) progression, and predict the emergence of cells in novel intermediate states of differentiation consistent with immunophenotypic characterizations of a mouse model of AML. In addition, we demonstrate that assuming continuous cell state may result in different dynamics when compared with the predictions of classical discrete models, particularly in anti-cancer drug resistance. We characterize the cases when the continuum and discrete models yield different dynamical patterns in the emerging heterogeneity in response to chemotherapy, and study its implications in designing therapies.
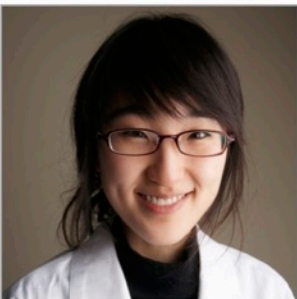
Biography:
Heyrim Cho received her B.S. degree in Applied Mathematics at KAIST, Korea in 2007 and her M.S. degree in Mathematics at the same institution in 2009. She received her Ph.D in the Division of Applied Mathematics at Brown University advised by Professor George Karniadakis. Her Ph.D research was on developing numerical schemes for stochastic modeling and simulations, particularly involving high-dimensionality. Currently, she is at the University of Maryland as a Brin postdoc fellow. Her research interest is in mathematical and computational biology, uncertainty quantification, and scientific computing.

Extreme Events:
Dynamics, Prediction and Mitigation

Mohammad Farazmand
Massachusetts Institute of Technology

A wide range of natural and engineering systems exhibit extreme events, i.e., spontaneous intermittent behavior manifested through sporadic bursts in the time series of their observables. Examples include ocean rogue waves, intermittency in turbulence, extreme weather patterns and epileptic seizure. Because of their undesirable impact on the system or the surrounding environment, the real-time prediction and mitigation of extreme events is of great interest. In this talk, I discuss some recent advances in the quantification and prediction of extreme events. In particular, I introduce a variational method that disentangles the mechanisms underpinning the formation of extreme events. This in turn enables the data-driven, real-time prediction of the extreme events. I demonstrate the application of this method with several examples including the prediction of ocean rogue waves and the intermittent energy dissipation bursts in turbulent fluid flows. I will also discuss future research directions in data-driven prediction and control of extreme events.

Biography:
Dr. Farazmand is a postdoctoral associate at Massachusetts Institute of Technology where he works on reduced-order modeling, prediction and control of extreme events in natural and engineering systems. He received his PhD in Mathematics from ETH Zurich in 2014 where he developed data-driven methods for extracting large-scale coherent structures from dynamical systems. Prior to joining MIT, he spent one year at Georgia Tech as the J. Ford Postdoctoral Fellow where he developed a new optimization method for discovering invariant solutions of partial differential equations. Farazmand received his BSc in Mechanical Engineering from Sharif University of Technology (Iran) and a MSc in Computational Engineering and Science from McMaster University (Canada).

Solving Linear Equations
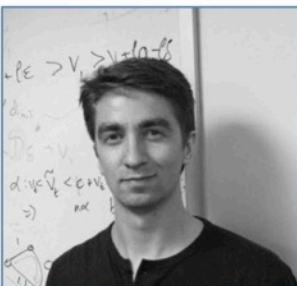
Rasmus Kyng
Harvard University

Solving systems of linear equations is one of the most fundamental algorithmic problems. In the past fifteen years, theoretical computer scientists have made tremendous progress in developing provably correct, asymptotically fast algorithms for solving structured linear equations.

I will present my 'Approximate Gaussian Elimination' algorithm, a very simple procedure for solving an important class of linear equations known as Laplacians. This algorithm may finally give us a practical, fast, and robust solver for these systems, making theoretical developments in the area relevant in practice. We will see some experimental evidence for this.

I will survey my research developing the asymptotically fastest known algorithms for solving linear equations in families known as Laplacians, Connection Laplacians, Directed Laplacians, and 3D truss stiffness matrices. These linear equations are used in algorithms for numerous problems in optimization, engineering, statistics, and data science broadly.

Next, I will discuss limits on fast algorithms for solving structured linear equations. My work on complexity lower bounds for solving structured linear equations shows that several classes of linear equations that seem only slightly more general than Laplacians are in fact as hard to solve as arbitrary linear equations.

Finally, we will see some recent work, which shows that ideas from linear equation solving and numerical analysis can be generalized beyond this setting to obtain new state-of-the-art algorithms for non-linear regression problems.

Biography:
Rasmus Kyng is currently a Postdoctoral Fellow at the Harvard Theory of Computation Group, mentored by Jelani Nelson. Before this, he was a Postdoctoral Research Fellow at the Simons Institute for the Theory of Computing at UC Berkeley. Rasmus obtained his PhD in Computer Science from Yale University under the supervision of Daniel Spielman. Rasmus has worked on algorithms for solving structured systems of linear equations, new approaches to convex optimization, learning and regression problems on graphs, random matrix theory, and fine-grained complexity lower bounds. His work won the FOCS 2017 Machtey award for best student paper.

Classical Verification of Quantum Computations

Urmila Mahadev
UC Berkeley

How can quantum computers be classically tested? This challenging question is interesting as a novel question about interactive proofs, as a practical question about the testing of near-term quantum devices, and as a philosophical question about the testing of quantum mechanics in the limit of high complexity.

In this talk, I will show that classical cryptography provides an elegant solution to this question: I will show that it is possible to classically verify quantum computations through interaction by relying on the assumption that quantum machines cannot break the cryptographic problem of learning with errors. This is achieved by constructing a commitment protocol in which a classical string serves as a commitment to an exponentially complex quantum state.

This talk will not assume prior knowledge of quantum computing or cryptography.

Biography:
Urmila Mahadev graduated in May 2018 with a Ph.D. in computer science from the University of California, Berkeley and continued as a postdoc. She is interested in quantum computation, complexity theory and cryptography.

An Aggregative Game Framework for the Analysis of Socio-Technical Systems

Francesca Parise
Massachusetts Institute of Technology

Many of today's most promising technological systems involve large numbers of autonomous agents that influence each other and make strategic decisions within a given infrastructure. Examples include demand-response methods in energy markets, opinion dynamics and targeted marketing in social networks, routing decisions in transportation systems or economic exchange and international trade in financial networks. The analysis of agents behavior and equilibrium outcome in these large scale systems necessitates the development of new theoretical and algorithmic tools that combine ideas from game, network and control theory.

In this talk, I discuss how aggregative games can help us achieve such a goal by providing a systematic framework for the modeling and control of large scale socio-technical systems. Specifically, I will touch on three models. First, I will present how "*average aggregative games*" can be used to model systems where each agent is affected by the aggregated actions of the rest of the population and how iteratively broadcast information can be used to coordinate agents behavior. Second, I will consider systems where agents interactions are heterogeneous and can be described by a network. I will present a variational inequality framework for the analysis of such "*network games*" which allows us to extend previous literature results, gain a systematic understanding of how network interactions affect the equilibrium outcome and plan targeted interventions based on agents centrality measures in social and financial networks. Finally, I will focus on a new game theoretical framework that I am developing to model strategic interactions in very large scale networks by using the concept of "*graphon games*" and I will illustrate how this framework can be exploited to design interventions that are robust to stochastic network variations.

Biography:
Francesca Parise is a postdoctoral researcher at the Laboratory for Information and Decision Systems at MIT. She defended her PhD at the Automatic Control Laboratory, ETH Zurich, Switzerland in 2016 and she received the B.Sc. And M.Sc. degrees in Information and Automation Engineering in 2010 and 2012,respectively, from the University of Padova, Italy, where she simultaneously attended the Galilean School of Excellence.

Francesca's main research interest is in control, network and game theory. She has worked on a broad set of topics, including systems biology, reachability analysis, distributed multi-agent systems, network analysis, aggregative games and opinion dynamics.

Francesca was recognized as an EECS rising star in 2017 and is the recipient of the Guglielmo Marin Award from the "Istituto Veneto di Scienze, Lettere ed Arti," the SNSF Early Postdoc Fellowship, the SNSF Advanced Postdoc Fellowship and the ETH Medal for her doctoral work.

Statistical Estimation Under Algebraic Constraints: Total Positivity

Elina Robeva
Massachusetts Institute of Technology

Nonparametric density estimation is a challenging statistical problem -- in general the maximum likelihood estimate (MLE) does not even exist! While the MLE is known to exist under certain shape constraints such as log-concavity, statistical rates are still exponentially slow in the dimensionality of the data. In this talk we introduce total positivity, a very strong form of positive dependence, as a constraint for density estimation, and we conjecture that it can alleviate the curse of dimensionality. Though they possess very special structure, totally positive random variables are quite common in real world data and exhibit appealing mathematical properties. Given i.i.d. samples from a totally positive distribution, we prove that the MLE exists with probability one (for more than 3 samples). We characterize the domain of the MLE, and give algorithms to compute it. If the observations are 2-dimensional or binary, we show that the logarithm of the MLE is a piecewise linear function and can be computed via a certain convex program. Finally, we prove statistical guarantees for the convergence of the MLE in two dimensions. This highlights one vein of my broader work on statistical estimation under algebraic constraints.

Biography:
Elina Robeva is currently a Statistics Instructor and an NSF Postdoctoral Fellow in Mathematics at MIT. After winning two Silver Medals for Bulgaria at the International Mathematical Olympiad, Elina went to Stanford where she earned a Bachelors degree with Honors in Mathematics and received several awards including the Dean's Award of Academic Excellence. For her PhD, Elina went to UC Berkeley and worked under the supervision of Bernd Sturmfels. Her doctoral thesis "Decomposing Matrices, Tensors, and Images" won the Bernard Friedman Memorial Prize in Applied Mathematics.

Discrete Differential Geometry and Its Applications

Andrew Sageman-Furnas
Technical University of Berlin

We are surrounded by curved surfaces, from unstructured polymer cell membranes and periodically woven fabric to large architectural structures, that are built from discrete units.

In this talk, I will discuss my recent results on the foundations of discrete differential geometry. This work unifies previous special classes of integrable discrete surfaces into a general geometric theory. Harnessing a similar perspective can illuminate aspects of the applications mentioned above. In particular, I will highlight recent collaborations that studied the role of geometry and elasticity in structured materials, and the role of expected network properties in determining force distributions under loading for unstructured materials.

Biography:
Andy Sageman-Furnas's research focuses on foundations and applications in the emerging field of discrete differential geometry (DDG), which develops analogues that preserve or reveal essential features of smooth geometric objects. His emphasis is on discrete curves and surfaces, and his work builds new theories with notions of curvature, unifying existing theories that preserve relationships to nonlinear PDEs and integrable systems. These ideas are not developed in isolation. In addition to linking smooth formulations of physical systems and computation, they provide insight into the underlying geometry of real-world materials that are inherently discrete, ranging from large architectural structures to woven wire mesh and molecular polymer networks.

A Modern Maximum-Likelihood Approach for High-Dimensional Logistic Regression

Pragya Sur
Stanford University

Logistic regression is arguably the most widely used and studied non-linear model in statistics. Classical maximum-likelihood theory based statistical inference is ubiquitous in this context. This theory hinges on well-known fundamental results---(1) the maximum-likelihood-estimate (MLE) is asymptotically unbiased and normally distributed, (2) its variability can be quantified via the inverse Fisher information, and (3) the likelihood-ratio-test (LRT) is asymptotically a Chi-Squared. In this talk, I will show that in the common modern setting where the number of features and the sample size are both large and comparable, classical results are far from accurate. In fact, (1) the MLE is biased, (2) its variability is far greater than classical results, and (3) the LRT is not distributed as a Chi-Square. Consequently, p-values obtained based on classical theory are completely invalid in high dimensions. In turn, I will propose a new theory that characterizes the asymptotic behavior of both the MLE and the LRT under some assumptions on the covariate distribution, in a high-dimensional setting. Empirical evidence demonstrates that this asymptotic theory provides accurate inference in finite samples. Practical implementation of these results necessitates the estimation of a single scalar, the overall signal strength, and I will propose a procedure for estimating this parameter precisely. This is based on joint work with Emmanuel Candes and Yuxin Chen.

Biography:
Pragya Sur is a fifth year Ph.D. student in the Dept. of Statistics at Stanford University, advised by Prof. Emmanuel Candes. Prior to joining Stanford, she received a Bachelor of Statistics in 2012 and a Masters of Statistics in 2014 from the Indian Statistical Institute, Kolkata. She is broadly interested in developing theory and methodology for accurate inference and prediction in high-dimensional models commonly used to analyze modern large-scale datasets. In parallel, she is interested in controlled variable selection and its connections to causality, and questions arising in the context of fair machine learning. She is a recipient of a Ric Weiland Graduate Fellowship in the Humanities and Sciences from Stanford University.

The Cost of Crushing:
Curvature-Driven Wrinkling of Thin Elastic Shells

Ian Tobasco
University of Michigan

How much energy does it take to stamp a thin elastic shell flat? Motivated by recent experiments on wrinkling patterns formed by thin floating shells, we develop a rigorous method (via Gamma-convergence) for evaluating the cost of crushing to leading order in the shell's thickness and other small parameters. The observed patterns involve regions of well-defined wrinkling alongside totally disordered regions in which no single direction of wrinkling is preferred. Our goal is to explain the appearance of such "wrinkling domains." Our analysis proves that energetically optimal patterns maximize their projected planar area subject to a shortness constraint. This purely geometric variational problem turns out to be explicitly solvable in many cases of interest, and a strikingly simple scheme for predicting wrinkle patterns results. We demonstrate our methods with concrete examples and offer comparisons with simulation and experiment.

This talk will be mathematically self-contained, not assuming prior background in elasticity or calculus of variations.

Biography:
Ian Tobasco is a James Van Loo Postdoctoral Fellow and Assistant Professor in the University of Michigan Department of Mathematics. He received a B.S.E. in Aerospace Engineering from the University of Michigan in 2011, and a Ph.D. in Mathematics from the Courant Institute of Mathematical Sciences at New York University in 2016.

Ian's research focuses on the Calculus of Variations and nonlinear Partial Differential Equations. He enjoys working on problems that sit at the interface between mathematics, physics, and engineering, where advances in pure mathematical analysis of the underlying equations can lead to scientific breakthroughs in the lab and vice versa. Ian's recent work involves the use of energy minimization principles to explain and classify the zoo of wrinkling, crumpling, and folding patterns exhibited by highly deformable, thin elastic sheets. His other interests include the design of optimal heat transport mechanisms in fluid dynamics, and their comparison with naturally occurring transport by turbulent flows, as well as the variational analysis of spin glasses.

Large-Scale Optimal Transport: Statistics and Computation

Jonathan Weed
Massachusetts Institute of Technology

Optimal transport is a concept from probability which has recently seen an explosion of interest in machine learning and statistics as a tool for analyzing high-dimensional data. However, the key obstacle in using optimal transport in practice has been its high statistical and computational cost. In this talk, we show how different notions of regularization can lead to better statistical rates—beating the curse of dimensionality—and state-of-the-art algorithms.

Biography:
Jonathan Weed is a fifth-year PhD student in Mathematics and Statistics at Massachusetts Institute of Technology. Prior to MIT, he received his AB in Mathematics from Princeton University. He researches theoretical aspects of machine learning, with a particular focus on the statistical analysis of data with geometric structure. He is a recipient of an NSF Graduate Student Fellowship, an MIT Presidential Fellowship, and the Josephine de Kármán fellowship.